LAD: A Design Model for Interactive Lifelog Applications

Aaron Duane[1], Hyowon Lee[2], Björn Þór Jónsson[1], and Cathal Gurrin[2]

[1]IT University, Copenhagen, Denmark
[2]Dublin City University, Dublin, Ireland

2022

Abstract

In this work we present the Lifelog Application Design (LAD) model and discuss some of the modern challenges associated with total-capture lifelogging. The LAD model is intended to address many of these challenges and support the design of novel interactive lifelog prototypes that target a broader range of application use cases. By remaining impartial to the specific lifelog data, technology platform, and application criteria, the model is structured in a way to have broad-ranging utility within the domain. We demonstrate the LAD model's utility by exploring two case studies and a retrospective analysis of VRLE, a real-world application prototype developed to examine the potential of lifelog retrieval in virtual reality. This work is based on the accumulation of insights garnered from involvement in a number of collaborative lifelog projects over the past decade. It is our goal to encourage future researchers to utilise the LAD model to support the design and development of their lifelog application prototypes and further solidify the model's contribution to the domain as a whole.

Keywords: lifelogging, human-computer interaction, multimedia analysis, interface design, user experience

1 Introduction

In an era already defined by personal data, the concept of lifelogging may at first seem redundant. Nearly every mainstream application we interact with captures or requests some form of personal data, ranging from basic information like name and age, to personal multimedia data in the form of images and videos. Yet lifelogging can be interpreted as distinct from the simple act of recording personal data. It is the deliberate intent to capture a unified digital record of

the totality of an individual's life experience [1]. This type of high-fidelity record could have broad-ranging application and benefit that yet remains unrealised.

Already supporting this endeavour, are tools and sensors that have been developed over the years to record a plethora of data as continuously, comprehensively and automatically as possible. However, despite this progress, applications designed to exploit this wealth of lifelog data are not in mainstream use, and even in research they rarely exit the prototype phase [2]. This is possibly due to a range of factors, but the most commonly discussed are concerns over privacy and the commitment required to continuously record one's life, as exemplified by Gordon Bell's MyLifeBits prototype [3]. Yet there is nothing preventing lifelog applications from prioritising privacy for its users and commitment to begin capturing life experiences will only come when demonstrable benefit can be presented to a mainstream audience. To date, lifelog application research has predominately focused on technical challenges [2] and the evaluation of non-interactive elements, such as novel algorithms to improve content analysis, and not on the elements which should directly support the user's interaction with the data [4]. To support such interaction, we require a more thoughtful approach to lifelog application research that is design-oriented and encourages a wider variety of prototypes to be developed and evaluated.

Since it is not feasible for every research project to recruit an expert in human-computer interaction to consult on their lifelog application development, the benefit of informed application design, which considers the intrinsic relationship between the lifelog data, the target access mechanism, and the application's core function, cannot be understated. To support this endeavour, we propose the Lifelog Application Design (LAD) model, an abstracted representation which is intended to inform the design of lifelog applications while remaining impartial to the specific data, platform, or use case being targeted. The LAD model is intended to help designers and developers of lifelog applications to more systematically investigate possible design choices by offering structured concepts on issues characteristic of the domain. Though the model is difficult to evaluate due to its generalised approach, it is based on the accumulation of insights garnered from an involvement in a number of collaborative lifelog-related projects within the community over the past decade. We will demonstrate the LAD model's practical utility by presenting a selection of lifelog application case studies, both theoretical and real-world, and how their design might be informed by following the guidelines proposed in this work.

2 Background

Though we have already defined lifelogging as a vision to capture the totality of a human's life experience [5, 6, 7, 1], more specificity can be useful as different researchers can often have diverging classifications. For example, some researchers prefer to avoid a *total capture* approach to lifelogging as they postulate that we remain wholly incapable of capturing the totality of human experience and therefore its utility remains difficult to define. These researchers

instead take a *situation-specific* approach to lifelog capture, where they collect targeted aspects of a person's life such as diet, mood, or physical performance, in a domain-focused effort to address specific use cases or research questions [8, 9, 10, 11, 12, 13]. Yet despite this, the pursuit of capturing the totality of human experience has had a notably positive impact on research to develop novel techniques for collecting and analysing personal data [14, 15, 16] especially visual data which has become one of the most popular methodologies to support the total capture of life experience. Further defining the utility of a total capture approach is a direct stimulus for the conception of the LAD model as it intends to encourage and support more robust user-centred lifelog prototypes to better evaluate the utility of such applications.

One shared aspect between these two approaches to lifelog capture is the generation of large volumes of data [16]. Visually accessing very large datasets to garner actionable insight is not a new phenomenon and there has already been research in the design community regarding different visualisation practices and techniques [17, 18]. However, visually accessing large-scale lifelog datasets based on total capture remains a very open area of research [19]. To date, the most common method of approaching total capture is through the continuous recording of images via a wearable camera, often worn around a person's neck. These cameras can passively capture several thousand images per day without any input from the wearer and will very quickly generate enormous datasets that become far too unwieldy for anyone to explore manually. To address this issue, researchers have relied heavily on automatic content analysis to organise and annotate this type of lifelog data to better support search and exploration [20, 21]. However, without an appropriate interface for users to efficiently interact with this additional metadata, the lifelog dataset will often remain just as difficult to analyse or explore.

This reminds us again how important thoughtful interaction design is when developing a lifelog application, and this extends not only to the application itself, but also to the target access mechanism and how it can impact the presentation of lifelog data. Researchers have explored the potential of common hardware platforms such as laptops, tablets and phones [22, 23], and there has even been research into other less conventional platforms such as virtual reality [24, 25]. It is clear there is motivation to explore lifelog application design using different technologies and in different contexts but there remains no clear blueprint on how best to proceed with this within the domain. One driving factor in recent years is the rising popularity of the Lifelog Search Challenge (LSC) [26], a real-time lifelog retrieval competition held at the ACM ICMR conference intended to serve as an evaluation criteria for the state-of-the-art in lifelog application research. Participants are invited to solve lifelog retrieval tasks over shared datasets [19] and are graded in terms of their speed, accuracy, and precision [4]. Though retrieval is an important aspect to lifelog analysis, there remain several other aspects which often go overlooked due to challenges in their evaluation. This is another area the LAD model seeks to address, which we explore in further detail in our next section.

3 Design Model

The LAD model is driven by three primary components which correspond to the three most variable domains of lifelog application design, namely the lifelog data itself, the target technology platform, and the application's overall design criteria. These components are summarised under the headings *Data*, *Technology* and *Criteria* respectively. It is important to recognise that the utility of the LAD model is not found in a specific combination of knowledge regarding these three domains, but rather, in understanding the relationship between how each domain should interact with one another. This concept is visualised in Figure 1 where we can observe the relationship between each domain and how the output of the model exists where these three bodies of knowledge intersect. In the following, we introduce each of these component domains in more detail to better understand their roles.

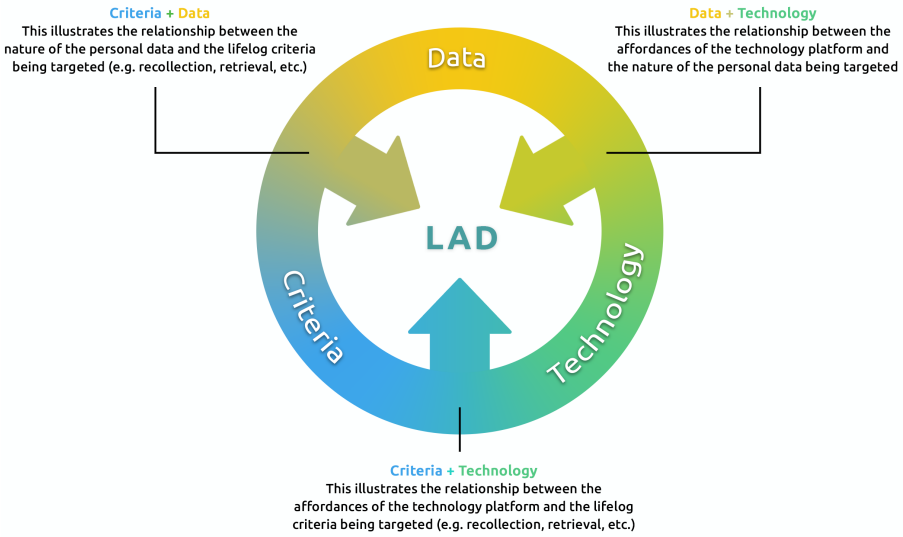

Figure 1: This figure visualises the implied understanding that emerges from the interactions between the target personal data, the technology platform, and the intended lifelog criteria within the LAD model

3.1 Domain Components

3.1.1 Data

The *Data* component of the LAD model might seem self-explanatory but specificity in this context is important when the category of information is inherently broad. Because lifelog data can theoretically encompass any type of personal media, the datasets are typically characterised by data which is collected auto-

matically, autonomously, and comprehensively [19] and results in multifaceted datasets with unique characteristics that require similarly unique retrieval techniques to explore effectively.

To date, the most common lifelog datasets released within the research community to encourage reproducible evaluations have consisted of a large corpus of images captured from the first-person perspective of a volunteer lifelogger [26]. These images are typically supplemented by other forms of personal data such as GPS coordinates, heart-rate, electrodermal activity, etc. However, in addition to this, there is also a significant amount of metadata which is generated via advanced content analysis. This analysis utilises techniques in computer vision to automatically tag images with concept descriptors, segment visually similar images into semantic events, and produce a host of other metadata which is continually being iterated upon.

Furthermore, lifelog datasets are rarely comparable, often containing data from multiple sources and sensors, or utilising completely different methods of analysis. As a result, it is very difficult to establish a unified strategy to explore lifelog datasets that is both versatile and effective. It is no surprise that data typically receives the most attention within the lifelogging community as it is perceived as the most valuable asset in this area of research [16]. Yet without proper consideration for the technology being used to interface with it and the specific design criteria being targeted, the true value of lifelog data can often go unrealised.

3.1.2 Technology

When we refer to *Technology* within the LAD model, we are referring to the target access mechanism or hardware platform which the lifelog application will be designed for. Within the research community this has most often been a desktop or laptop computer paired with a mouse and/or keyboard, however, there has been some effort at exploring other technologies [22, 23, 27] making use of the increasing body of design knowledge for non-desktop interaction platforms [28]. Yet despite the prevalence of more conventional platforms, it is important to note that any technology could conceivably be utilised in the design and development of a lifelog application. The limitation should not be on the technology itself, but rather its relationship to the lifelog data, and the design criteria the researchers intend to pursue.

For example, one might not initially consider a fridge as a viable platform for a lifelog application, but if one considers the advent of smart fridges equipped with advanced sensors and a touchscreen interface, this perception might be quickly reconsidered. Furthermore, it is important not to conflate the technology platform as a whole with its primary human interface device. For example in the context of our fridge, one might consider the technology comparable to that of a touchscreen tablet since the user interacts with the fridge via a similar touchscreen. Though these two technologies share similarities in their primary human interface mechanisms, they should not be considered as comparable technologies within the LAD model. There are numerous important

distinctions which must be considered, for example that a tablet is portable whereas a fridge is highly static.

More important, is how users interact with these separate technologies. A user may only interact with a fridge a handful of times a day, and for very predictable reasons. This means any effective lifelog application should accommodate for this behaviour and how it might impact the other primary components of the LAD model. This is also true if there is intent to target more than one type of access mechanism for a single lifelog application. It is then necessary to recognise what distinguishes the different technologies, and how this may impact the relationship between the other components of the LAD model as well. If the technologies are sufficiently different, it may warrant the addition of very different features, or even a complete redesign of the entire application as a whole.

3.1.3 Criteria

The *Criteria* component of the LAD model can often be overlooked in lifelog application design and yet it is fundamental in the design of an effective exploration tool. When we refer to lifelogging criteria within the LAD model, we are referring to the design principles implemented by the application to support specific use cases that serve a practical benefit of lifelogging. A list of such criteria was first proposed by Sellen and Whittaker [2] as a set of potential benefits a lifelog application could provide for its users. The benefits are categorized using the researchers' mnemonic five R's: recollection, reminiscence, reflection, retrieval and remembering intentions. We will describe these five R's and their relationship to the other components of the LAD model in the next section.

It should be noted that the five R's are not necessarily an exhaustive list of criteria, but by utilising these initial lifelog benefits, we can begin to define guiding principles for practical lifelog application design. It is also worth noting that these criteria only apply to lifelog applications targeting datasets based on total capture, rather than situation-specific capture. For applications targeting selective lifelog data, the design criteria are inherently defined by the domain of interest being captured. For example if researchers were focused on an individual's health and well-being, they may capture data such as heart-rate, physical activity, or galvanic skin response, and have specific research questions which become the natural focus of the lifelog application. In contrast, when we are generating datasets determined to capture the totality of life experience, we need to be more thoughtful in defining the specific criteria being targeted by a lifelog application to maximise its effectiveness as a platform for analysis.

Since lifelogging is not currently a ubiquitously practiced activity, we are likely to see more new and specific criteria appear in the coming years. As such, there is room for more systematic analysis in order to create such novel lifelog-related application areas (either based on the original five R's or outside of them). Identifying users' unmet but inherent needs will likely require several iterations of conceptualisation, design, prototyping, and usability testing (c.f. behavioral design in [29]). Similar to how many major innovations in the history

of technology have materialised, only by first seeing and using the tangible applications will we be able to know there was a new need which could be satisfied [30]. The LAD model is intended to help this systematic exploration by offering the component concepts and scope to start considering the invention of novel applications in the lifelogging domain.

3.2 Inter-Domain Component Relationships

Now that the individual components of the LAD model have been introduced, we will discuss how these three bodies of knowledge can interact with one another within the model itself, as illustrated in Figure 1.

3.2.1 Data and Criteria

Previously we introduced the five R's which represent the potential benefits a lifelog application might provide for its users and which we define as our initial list of primary criteria which should be accommodated for within an effective lifelog application. To explore how the *Criteria* component interacts with the *Data* component of the LAD model, we must first describe how each of these five R's are defined, and then discuss how they might interact with an application's target *Data*.

Recollection is described as the ability of a lifelog to support the simple act of recollecting, or reliving a specific life experience. This involves thinking back in detail on a past experience, sometimes referred to as an episodic memory [31, 32]. This type of remembrance enables us to mentally retrace our steps, useful for a host of practical purposes such as the locating of lost property or remembering faces and names by recalling when and where we met someone. Many lifelog applications implicitly address recollection, or remembering personal experiences [33], and there is well-known psychological literature that there is a strong connection between autobiographical memories and visual images [34]. In the context of the LAD model, this implies lifelog applications seeking to target recollection should try and focus on visual data such as images or videos in their core design, and to more carefully consider the special characteristics of this visual data.

Reminiscence can be observed as a specialised form of recollection, as it is still a form of reliving one's life experiences, but specifically for the purpose of emotional or sentimental reasoning as opposed to practicality; for example, watching a home movie of one's wedding or flipping through a photo album of a newborn child. If researchers want to support reminiscence, visual data can still be very beneficial but other factors also become more relevant, such as optimising the sharing of the data with others. In addition, it may be more appropriate to emphasise the temporal aspect of data presentation, such as highlighting the time of the day (morning, afternoon, etc.) or drawing attention to a specific amount of time which has passed between two or more memorable events.

Reflection refers to the ability of a lifelog application to support more abstract representations of personal data to facilitate reflection on, and reviewing of, past experiences; for example, examining patterns in one's behaviour over time, which could provide useful information about physical activity or emotional states in different contexts. This could then be related to other data about health or well-being. Reflection might also involve looking at one's past experiences from different angles or broader perspectives where the value is not in reliving past experiences but in seeing things anew and framing the past differently. Within the LAD model, this implies lifelog applications seeking to support reflection should focus more on abstraction, offering flexible and novel methods for viewing personal data in ways that might surprise or educate its users. Emphasising data visualisations which efficiently convey such abstractions will need to be especially cognisant of the temporal aspect of the data.

Retrieval refers to the ability of a lifelog application to support the retrieval of specific digital information the user has encountered such as images, videos, documents, etc. In this sense, retrieval can incorporate elements of the previous lifelog design criteria, such as how the retrieval of a specific email might support recollection, or the retrieval of a specific image might support reminiscence. Retrieval can often depend on inferential reasoning, such as trying to deduce keywords in a document or thinking about the document's other likely properties. The consideration of information properties need not involve recollection of past experiences at all as long as other ways are available for finding the desired information. In the context of the LAD model, lifelog applications seeking to support retrieval should focus on efficient ways of searching through large heterogeneous collections of data and provide access to metadata that might support more effective filtering, ranking and search. Within the lifelogging community, retrieval is the most commonly evaluated aspect of a lifelog system due to the relative ease of evaluating a criteria which does not necessitate the owner of the dataset being the user of the system (although in some more general usage scenarios the searcher might indeed need to be the owner of the dataset).

Remembering Intentions differs from our previous lifelog design criteria in that it does not relate to retrospective memories, but rather to prospective memories, or remembering things you intend to do. For example, this could be remembering to run errands, take medication, or show up for an appointment. To support this within the LAD model, designers would need to focus on delivering timely cues in appropriate contexts if they are to provide effective reminders. The nature of prospective events means the lifelog application would likely utilise some form of calendar or diary which would need to integrate seamlessly with the target data.

We have now introduced an initial list of primary lifelog criteria, and discussed their respective relationships with various data. However, it is important to remember that the relationship between *Criteria* and *Data* operates equally in both directions. The LAD model does not necessitate a specific criteria to be selected before targeting specific data. It is equally viable to begin with target data, and then apply the model to select the most relevant criteria.

3.2.2 Technology and Data

The relationship between lifelog data and the technology being utilised to analyse it is often implicitly acknowledged within the research community rather than explicitly defined. This is because the most common technology used as a platform for lifelog applications is a desktop or laptop computer, or more precisely, a screen paired with a keyboard and mouse or trackpad. Though there is an enormous variety of applications developed for this type of technology, interactive design has matured to the the point that there are fundamental interaction elements which can often be observed across all such applications (e.g. point and click, scroll-bars, input boxes, drop-down menus, etc.). Lifelog application prototypes are not exempt from this, and as such, many design decisions in the development of a lifelog application often go unquestioned as the researchers simply follow the trend of what has worked previously. This, in and of itself, is not bad practice, but complications arise when we begin to consider other more divergent technologies, especially those which differ drastically from that of a conventional desktop computer.

There can be a tendency when developing for a new technology to simply adapt application elements from pre-existing platforms. This is convenient for obvious reasons as it significantly reduces development overhead and allows designers and developers to reuse assets and code logic. And in some cases this type of approach can have varying degrees of success, for example an application developed for a touchscreen tablet might be easily adapted to work on a touchscreen smartphone with the potential for minimal loss in the user experience. However, if we were to adapt that same lifelog application for virtual reality, it would be far more effective to consider the specific affordances the technology provides and how it relates to the lifelog data being targeted, rather than trying to simply augment previous design strategies for a virtual setting. As a notable example, the key affordance of virtual reality platforms is the addition of a third spatial dimension and six-degrees of freedom within the virtual space. Understanding these special characteristics of the technology, and then building a suitable lifelog application that leverages its affordances according to those characteristics, is likely to more optimally support analysis of the target data.

It is also important to remember again that the relationship between these components of the LAD model operate in both directions. Just as we should consider what affordances a specific technology provides, we must also consider the type of lifelog data being targeted and how it could be best utilised. For example, if the data is primarily visual, such as images, it would be beneficial to target a technology which affords sufficient visual acuity, such as a widescreen monitor or television, rather than a smartphone. Or, in the event that a smartphone is a design requirement, then it is necessary to consider how the design can be adapted to account for any potential loss in visual acuity. The relationship between *Data* and *Technology* within the LAD model underlines these considerations and highlights their significance in the development of an effective lifelog application.

3.2.3 Criteria and Technology

To convey how the *Criteria* and *Technology* components of the LAD model interact with one another, we will again utilise our initial list of primary lifelog criteria inspired by the 5 R's. It warrants repeating that, as before, the relationship between these components operates in both directions, and therefore the decision to approach it from the perspective of *Criteria* is merely demonstrative.

If we recall that *recollection* as a criteria corresponds to the simple act of reliving one's memories or life experiences, the LAD model should direct us to consider any technology which effectively supports the psychological cues necessary to evoke this. Since we have already established there is prior evidence to suggest visual cues are especially effective in this context, we should favour technologies that provide the best affordances for visual presentation. However, we must also consider the context within which the user is intended to relive these experiences. For example, it may be tempting to target a widescreen smart TV due its large screen size and perceived relationship to visual media. However, if the user is intending to engage in *recollection* for a more direct purpose, such as recalling where they left their keys, this would make the television inconvenient due to its static nature, and therefore it might be better to consider a sufficiently high-resolution smartphone as the trade-off between screen size is offset by its mobility in this scenario.

As *reminiscence* is closely connected to the criteria of *recollection*, some of the same inferences might be concluded in choosing an appropriate technology. However, once again we should not neglect to consider the subtle shift in focus. For *reminiscence*, the goal is to relive memories for meaningful and sentimental purposes. This may suggest a technology better suited to a more comfortable or slow-paced environment. This might lead us to consider again a smart TV, or in seeking to leverage a highly immersive visual experience, we might even consider a virtual reality technology.

The focus of *reflection* as a criteria is on identifying patterns in our behaviour via more abstract representations of life experiences. In this context, the LAD model might encourage us to consider technologies which most effectively display and afford high-fidelity interactions with these representations. This may lead us to consider a conventional computer or sufficiently large tablet device, and perhaps discourage smartphones or small tablets due their reduced screen size which may negatively impact the presentation and precise interactions with abstract representations. Of course this does not preclude the possibility of such technologies as candidates, as their usage may be a potential prerequisite. This simply means that the utility of the LAD model is in highlighting potential problem areas one may need to be cognisant of, depending on the context of the application use case.

Since *retrieval* as a lifelog criteria relies heavily on the successful navigation of heterogeneous data collections and access to metadata that supports effective filtering, ranking, and search, the LAD model would suggest targeting a technology that adequately supports the user's interaction in this context. For example, if we consider a lifelog application which generates retrieval queries

based on generic user input, the technology should effectively support these inputs. This could be as simple as efficient text entry for search, the effectiveness by which they can navigate ranked results, or the general ease by which any metadata can be explored. Due to the ubiquity of personal computers which utilise recognisable human interface devices like physical keyboards and mice, they may be an obvious initial consideration. Yet smartphone and tablet devices now often come equipped with sophisticated touchscreens and supplementary accessories, so their viability can be equally considered depending on the specific requirements of the application or research.

When we consider the criteria of *remembering intentions* and its unique reliance on prospective, rather than retrospective, memories, the impetus to rely on prompt and contextual reminders becomes evident. Since the goal is to help the user remember things that they intended to do, one might consider a lifelog application that relies on some kind of calendar or diary notifications. Framed in the context of the LAD model, this might initially suggest a smartphone as a suitable technology due to its portability and likelihood of being in close proximity to the user for the most pertinent and contextual reminders. As before, this example is not intended to negate the possibility of other technologies, but merely to illustrate an effective candidate based on initial projections within the model.

3.2.4 Data, Technology and Criteria

Now that we have introduced each component and summarised their binary two-way connections within the domain-relationship model (Figure 1), we can begin to better explore how the LAD model is intended to exist within the confluence of these three bodies of knowledge. We achieve this by first exploring two theoretical case studies, before moving onto a real-world case study in the section following, where all three cases are framed in the context of the LAD model. These examinations are intended to represent the design and development of diverse lifelog application prototypes utilising the model to support and inform their respective approaches.

4 Theoretical Case Studies

It is important to remember that lifelog application development can begin with any component of the LAD model already partially or fully established. For example, it is not uncommon to already have collected or have access to a lifelog dataset and even have an intended criteria for its usage. In contrast, it is also possible that a project might already have an application in place on a specific technology platform, such as a smartphone, and want to expand their application to encompass a new lifelog dataset or criteria. The aim of the LAD model is to frame any previous work or constraint in a specific context to highlight the relevant domain relationships and thereby maximise the potential of any resulting lifelog application design. In the context of our two theoretical

cases studies, we have endeavoured to craft realistic scenarios to better convey the real-world utility the LAD model is intended to provide.

4.1 Case Study 1: Reminiscence using personal videos and a smart TV

In our first case study, the project already has access to a personal dataset of videos actively being generated by one individual. There are close to 1,000 videos to date and each video varies from a few seconds to several minutes in length. The content of these videos is characterised by the owner as ranging from mundane day-to-day activities to important life events, such as birthdays and family occasions. In addition to the video data, there is associated metadata for each video provided by the various cameras which were utilised during the data collection process. The owner of the dataset, henceforth referred to as the lifelogger, intends to continue adding to the dataset over time, and is interested in developing an application to explore it in a way that is meaningful to them.

Upon consulting the guidelines outlined in the LAD model, the researchers identify reminiscence as the primary *Criteria* this lifelog application will address. This is supported by the lifelogger's desire to explore the data in a way that is sentimental and emotionally engaging and is reinforced by the visual nature of the lifelog data which better lends itself to effective reminiscence due to the strong connection between autobiographical memories and visual media [34, 2]. There is no constraint on the target access mechanism, other than it should be readily accessible by the lifelogger.

Considering the *Data* and *Criteria* which has already been established, the researchers decide to target a smart television as their chosen *Technology*. This supports the primary constraint in that the lifelogger already owns a smart television, but also complies with the principles described by the LAD model in that the larger screen and increased pixel density will comfortably support the nature of the lifelog data being explored. Furthermore, the lean-back usage environment [35, 36, 37] of the television in a living room is congruent with the relaxed and sentimental aspects of the target lifelog *Criteria*. The researchers begin working on their prototype application, utilising a design framework and development workflow which is familiar to them and is appropriate for the size of their team. As the LAD model is intended to be agnostic to low-level design and development decisions, the details of this work are outside the scope of this case study.

Upon completion of an initial prototype, the researchers identify that the application could easily be adapted to support recollection in addition to its primary criteria. This was determined upon considering the overlap between certain features already implemented within the prototype and how these reflected the overlap between use cases for recollection and reminiscence, which were also both equally supported by the targeted data and access mechanism. One notable example was the implementation of a timeline which temporally visualised the video data captured by the lifelogger. This feature was very applicable to a common recollection use case of simply remembering what happened

on a specific day in the past. The day did not need to be linked to a special occasion for the lifelogger but rather served as a cue to recall their location at various points in their past.

4.2 Case Study 2: Reflection using biometric data and a smartphone

Our second theoretical case study is more unusual in that the researchers do not have access to any visual lifelog data such as images or videos. Instead, they possess some robust and comprehensive biometric datasets collected by three individuals which primarily consist of quantified-self [9] data, such as heart-rate, electrodermal activity, step count, and geolocation. This *Data* was collected daily and consistently over the course of one year but does contain occasional gaps where sensors failed due to low battery or malfunctioning software. The datasets also contain some other data types such as light intensity and blood sugar level but they were deemed to be less relevant as the data was captured too sporadically and the respective sensors were not always accurate. The data itself was collected as part of multiple separate projects and as a result does not share a consistent theme or focus and the lifeloggers are looking for new ways to utilise it in their research.

Upon reviewing the LAD model, it becomes apparent that the most appropriate *Criteria* the researchers could focus on is reflection, as the biometric data they have collected serves as a very effective means to examine patterns in one's behaviour over time. They also decide to target a smartphone as their primary *Technology* due to its ubiquity of use and partly due to the personal and discrete viewing the platform naturally affords in accessing such personal data. However, in framing their design decisions in the context of the LAD model, they come to appreciate that to effectively benefit from reflection on their biometric datasets, it will require an access mechanism that supports large and abstract visualisations which can be easily adjusted via numerous user interface controls. The researchers acknowledged that a laptop would be a preferable platform for this type of interaction, but decide that the trade-off in form factor and convenience by targeting a smartphone is acceptable, especially since they have more expertise in developing for this technology.

As in our previous case study, the researchers begin designing and developing their prototype application utilising a design and development strategy that they are already familiar with. Maintaining the target *Criteria* of reflection, they begin to implement features which support the framing of the lifelog data from different perspectives in order to identify patterns in behaviour. One implementation was the aggregation of heart-rate and electrodermal activity into a stress indicator which could be tracked temporally via a line graph. This would enable the lifeloggers to establish patterns of stress at various points in their lives. As development progressed, the researchers continue to identify constraints in relation to the smartphone *Technology* and their target *Criteria* of reflection. However, in remaining cognisant of these constraints, they were able to augment their design and alleviate the potential for any excessive cognitive

load [22] for their users.

The researchers eventually complete and evaluate several iterations of their lifelog application. During this time, the original lifeloggers have continued to collect biometric data but the successful implementation of their system using the LAD model has encouraged the researchers to collect new types of data to further enrich their analysis platform. One of the lifeloggers even volunteered to begin collecting images via a wearable camera as they were interested in finding out exactly what they were doing during specific points in the reflection of their data. This pivot in the variety of lifelog data being collected contributed to the final decision to develop an alternate version of the prototype, namely a web application, as much of the existing software could be reused in its development. This version of the prototype, which would utilise a larger screen and a mouse and/or keyboard, would enable more rigorous analysis of the lifelog data to support reflection, but also enable the comfortable viewing of the more visual lifelog data now intended to be collected. This decision also led to the inclusion of recollection as a secondary focus for the application due to the potential for utilising the visual lifelog data as a cue for remembering specific past experiences.

5 VRLE: Retrieval using egocentric images in virtual reality

We have now described two theoretical case studies with the intent to explore the broad-ranging application of the LAD model. This is important as our primary goal is to encourage a wider variety of lifelog application to be considered for active research. However, it would be beneficial to apply the LAD model to a real-world case study to demonstrate its utility more explicitly. To that end, we will describe the development process behind VRLE, a first-generation virtual reality lifelog application. We should emphasise that the LAD model did not inform nor support the original design or development of VRLE, as the prototype was developed prior to the model's inception. As such, this discussion will serve as a retrospective analysis framed in the context of the LAD model.

The VRLE (Virtual Reality Lifelog Explorer) application prototype began development in 2017 and was intended as a case study to explore the effectiveness of virtual reality as a basis for lifelog exploration. Similar to our theoretical case studies, this research project began with one component of the LAD model firmly established, in this case the *Data*. Due to numerous factors, most notably availability and privacy concerns, the only lifelog dataset of sufficient size and quality at the time was available via participation in the first Lifelog Search Challenge (LSC) workshop at ACM ICMR 2018 [26]. The LSC is an annual competition carried out during the ICMR conference where competing lifelog prototypes solve known-item search tasks in real-time. The 2018 dataset consisted of 27 days of visual data captured by 1 lifelogger which corresponded to 41,681 egocentric images taken in 72 different geographic locations [38]. In

addition to this, data such as biometrics and the lifelogger's multimedia habits were also made available, but this was not as comprehensive and therefore less emphasised. Since the research project was now utilising the LSC dataset, it was natural to target the LSC itself as an evaluation metric, where we could compare our prototype to the state-of-the-art. As such, in the context of the LAD model, our target *Criteria* became lifelog retrieval, as the LSC exclusively targeted tasks involving the retrieval of visual data captured by the lifelogger.

It should be noted at this point that a virtual reality platform was not initially considered as a target access mechanism for this research. However, the Oculus Rift and HTC Vive both had their commercial release in 2016 and several smartphone-based virtual reality applications, such as Google Cardboard and Samsung Gear, had become popular. The increased public awareness of virtual reality as a medium encouraged a reevaluation of its potential application. Research at the time suggested that the most valuable aspect of virtual reality was its highly immersive quality and the degree to which it projected stimuli onto the sensory receptors of users in a way that was extensive, interactive, and vivid [39]. There had also been research implying that actively using more of the human sensory capability and motor skills was known to increase understanding and learning [40] and that immersion could greatly improve user recall [41]. There was general belief motivating much of this research that the technology would lead to more natural and effective human-computer interfaces [42]. This gradual focusing of the virtual reality platform naturally became a primary aspect of the project due to its compelling impact on user interaction [43]. The HTC Vive was ultimately chosen as the target *Technology* due to its advanced features and higher fidelity at the time of purchase.

The consideration and eventual targeting of the HTC Vive are an initial example of where the LAD model would have been useful in framing design decisions by highlighting the impact choosing it would have on our target *Data* and *Criteria*. However, before we discuss this in further detail, we must first outline the final design of the VRLE prototype to provide additional context. The application can be conveniently divided into two primary components, with one component focusing on the formulation of user queries in the virtual space, and the other component focusing on navigating the visual data returned from those queries in that same space.

5.1 Querying Component

The querying component consisted primarily of a 3D user interface which appeared in the virtual environment as a 2D plane. To interact with this virtual interface, the user simply had to reach out with one of the HTC Vive's controllers and directly touch the relevant UI element. To aid with precision, the virtual representation of the controllers was outfitted with a long protuberance which we referred to as a drumstick due to its appearance (see Fig. 2). To prevent accidental interactions, a drumstick would only interact with the interface when the user held down the controller's trigger, which would cause the drumstick tip to turn white. The query interface was itself divided into two parts,

corresponding to the two types of metadata the user could filter by, tags and time. When we refer to tags, we are referring to concept descriptors produced by automatic content analysis. This process would label every image in the dataset with a set of words attempting to describe the image's content, such as coffee, vehicle, desk, etc. The time metadata is more intuitive, as it corresponds to the temporal aspect of the query, such as limiting the query to a specific month, day, hour, etc.

Figure 2: VRLE - 3D user interface for generating queries

The user could generate filter queries using any combination of the available metadata. For selecting tags, this required an additional layer of interaction as the user needed to be able to search through the available tags. This took the form of a search box where the user could begin typing using their drumsticks and a virtual keyboard (see Fig. 2). As the user typed each letter, matching tags would appear which they could then select to add to their query, or remove if already selected. Finally, the user could re-position the interface within the virtual space at any time to adjust to their preferred height and viewing angle.

5.2 Visualisation Component

Upon submitting a query, VRLE would begin loading the results of that query into the virtual space perpendicular to the direction the user was facing. These results appeared as a wall of images which could be scrolled left or right by clicking a button on either controller and performing a throwing motion in the preferred direction. The results were grouped temporally based on image similarity to produce what became referred to as lifelog events. This type of semantic grouping was necessary as the lifelogger would often capture several hundred near-identical images in a row, such as when they were working at a desk or watching something on television. Ranking and visually displaying all of the images within such events would quickly overwhelm and inhibit the user's ability to effectively navigate the results of their query. Depending on the length of an event, we selected up to 9 images arranged in a 3x3 grid to serve as a preview of the event's content (see Fig. 3). The images selected for these event previews were based on how many of the user's queried tags each image

contained. Above the event preview we included some additional metadata such as start and end time to further contextualise the event for the user.

Figure 3: VRLE - Scrollable list of ranked lifelog events

The events were ranked from left to right based on the number of images contained in the event which were associated with the user's queried tags. Though this approach to grouping and ranking lifelog images empowered users to more quickly identify relevant sets of results, it was still necessary to enable users to examine all of the images in an event when necessary. To achieve this, the user could point their VR controller at any image in an event preview; this would cause a contextual menu to appear alongside the controller and a solid blue beam to confirm what image the controller was pointing at. Using this contextual menu, the user could submit the target image and this would cause all the images in the event to appear in a timeline in front of its event preview (see Fig. 4). The submitted image from the event preview would then serve as a reference point for how far into the timeline of the event to scroll to, with earlier images appearing on the left and later images on the right.

The user could interact with an event timeline via a secondary context menu which worked identically to how the user interacted with the previous menu. However in this instance some additional metadata and contextual options are exposed. Most notably, the user is provided with the option to 'See Similar' images in relation to the image being pointed at (see Fig. 4). Selecting this option would reload the timeline with a set of images from the dataset which most closely resembled the target image but were not present in the current event timeline. This feature was intended for situations where the user had retrieved a relevant set of results but they were missing some minor visual detail.

Figure 4: VRLE - Scrollable image contents of one event

5.3 VRLE and the LAD model

Now that we have described the primary components of the VRLE application prototype, we can begin to reflect on some of its key design decisions in the context of the LAD model. As mentioned earlier, the first and most noteworthy decision was to select the HTC Vive as the application's target *Technology*. While this decision is congruent with the visual nature of the target *Data*, namely egocentric images, as it provides a broad range of utility in visualising this data in a virtual setting, its congruity with the target *Criteria* of retrieval is less apparent. This is because effectively retrieving relevant digital documents relies heavily on two main factors; the quality of the documents' metadata and the ability of the application to support queries using that metadata. These queries can vary in their complexity, but most often rely on a set of inputs which correspond to specific facets of the metadata which the user wishes to search, filter, or rank. Interacting with such inputs in a virtual environment remains a challenge as the user does not have access to familiar input devices such as a physical mouse and keyboard. While the precision of the mouse can be replicated using various virtual substitutes such as a laser pointer, the benefit of a physical keyboard is not so easily reproduced when the user cannot effectively utilise multiple fingers in the virtual world.

These limitations of the *Technology* in relation to the *Criteria* could have been alleviated by various means, for example, relying on voice recognition instead of typing, or replacing the VR controllers with a substitute that more closely exploited the dexterity of the human hands. However, without the LAD model to highlight these issues, or to potentially dissuade the use of the HTC Vive as a platform entirely, a bottleneck was quickly identified in the querying component of the VRLE prototype during testing. On average users took slightly more time to construct retrieval queries than they would normally using a conventional desktop alternative [44]. However, as the LAD model also suggested, this bottleneck was partially offset by the relationship between the *Technology* and *Data*. Users reported feeling more immersed interacting with the image data in virtual reality and on average found what they were looking for slightly faster when compared to the conventional alternative [44].

Overall the VRLE research suggested that retrieving lifelog images using the HTC Vive was not notably more effective than a conventional system, but also highlighted that it was not notably less effective either. This result is not surprising when we consider the relationship between the target *Criteria* and *Data*. As noted earlier, the retrieval of visual data relives heavily on the quality of its metadata which in turn most often relies on non-interactive content analysis. This suggests that two retrieval systems, even when using very different platforms, targeting the same metadata will most likely be comparable in terms of performance as long as both platforms can adequately support user querying and data visualisation. In essence, the fundamental relationships arising from two components of the LAD model (*Criteria* and *Data*) outnumber the relationships arising from just one component (*Technology*) and therefore restricts the potential impact that component can have on the lifelog application's design.

In a situation like VRLE where two of the three components of the LAD model were unable to be altered, the model itself still provides utility. For example, it highlights important relationships between core components so researchers can be cognisant of potential problem areas from the onset of the research life cycle. Furthermore, it could provide the necessary stimulus to shift the focus of the research at the start of the cycle when such changes are more viable. In the context of VRLE, one option may have been a shifting of the target *Criteria* to something like reminiscence instead of retrieval. This is because reminiscence does not rely as heavily on metadata and the focus on reliving one's life experiences for emotional and sentimental reasons could be strongly supported by virtual reality as a *Technology*. It is easy to imagine a virtual personal space where you could relive important memories without relying on high fidelity querying of metadata. Nevertheless, it should be noted that despite these observations, VRLE managed to win the first LSC competition, albeit by a small margin. Though there are likely extenuating circumstances arising from the relative infancy of the competition itself [26], this highlights the broader issue of lifelog applications lacking a robust process for effective design and comparison which the LAD model intends to support.

6 Conclusion

We have discussed some of the modern challenges associated with total capture lifelogging and highlighted the importance of thoughtful interaction design to encourage a more robust generation of lifelog application to be considered within the domain. To support this, we have introduced the Lifelog Application Design (LAD) model which is based on the accumulation of insights garnered from an involvement in a number of collaborative lifelog-related projects within the community over the past decade. By remaining impartial to the specific lifelog data, technology platform, and application criteria, the design model is intended to have broad-ranging utility in the development of new lifelog application prototypes. We have demonstrated this utility through an analysis of two theoretical case studies and a retrospective analysis of VRLE, a real-world

prototype developed to examine the potential of lifelog retrieval in virtual reality. It is our goal to encourage future researchers to utilise the LAD model to support the design and development of their lifelog application prototypes and further solidify the model's contribution to the domain.

References

[1] M. Dodge and R. Kitchin. "'Outlines of a World Coming into Existence': Pervasive Computing and the Ethics of Forgetting". In: *Environment and Planning B: Planning and Design* 34.3 (2007), pp. 431–445.

[2] A. Sellen and S. Whittaker. "Beyond Total Capture: A Constructive Critique of Lifelogging". In: *Communications of the ACM* 53.5 (2010), pp. 70–77.

[3] J. Gemmell, A. Aris, and R. Lueder. "Telling stories with MyLifeBits". In: *IEEE International Conference on Multimedia and Expo, ICME 2005.* Vol. 2005. 2005, pp. 1536–1539.

[4] D. Kelly. "Methods for Evaluating Interactive Information Retrieval Systems with Users". In: *Foundations and Trends in Information Retrieval* 3.2 (2009), pp. 9–14.

[5] S. Mann. "Wearable computing: A first step toward personal imaging". In: *Computer* 30.2 (1997), pp. 25–32.

[6] S. Mann. "Continuous lifelong capture of personal experience with EyeTap". In: *Proceedings of the the 1st ACM workshop on Continuous archival and retrieval of personal experiences - CARPE'04* (2004), pp. 1–21.

[7] S. Mann et al. "Realtime HDR (High Dynamic Range) Video for Eyetap Wearable Computers, FPGA-Based Seeing Aids, and Glasseyes (Eyetaps)". In: *25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE).* 2012.

[8] M. A. Barrett et al. "Big Data and Disease Prevention: From Quantified Self to Quantified Communities". In: *Big Data* 1.3 (2013), pp. 168–175.

[9] J. Meyer et al. "Beyond Quantified Self: Data For Wellbeing". In: *Proceedings of the Annual ACM Conference on Human Factors in Computing Systems, Extended Abstracts (CHI EA).* Toronto, ON, Canada, 2014, pp. 95–98.

[10] M. B. Hoy. "Personal Activity Trackers and the Quantified Self". In: *Medical Reference Services Quarterly* 35.1 (2016), pp. 94–100.

[11] S. A. Bahrainian and F. Crestani. "Are Conversation Logs Useful Sources for Generating Memory Cues for Recalling Past Memories?" In: 8 (2017), pages.

[12] S. A. Bahrainian and F. Crestani. "Towards the Next Generation of Personal Assistants". In: vol. 9. 2017, pp. 169–176.

[13] S. A. Bahrainian and F. Crestani. "Augmentation of Human Memory: Anticipating Topics that Continue in the Next Meeting". In: *ACM Reference*. 2018, pp. 150–159.

[14] M. Czerwinski et al. "Digital memories in an era of ubiquitous computing and abundant storage". In: *Communications of the ACM* 49.1 (2006), p. 44.

[15] C. Gurrin et al. *Exploring the Technical Challenges of Large-scale Lifelogging*. 2013.

[16] C. Gurrin, A. F. Smeaton, and A. R. Doherty. "LifeLogging: Personal Big Data". In: *Foundations and Trends in Information Retrieval* 8.1 (2014), pp. 1–125.

[17] L. Wang, W. Guanghui, and C. A. Alexander. "Big Data and Visualization: Methods, Challenges and Technology Progress". In: *Digital Technologies* 1.1 (2015), pp. 33–38.

[18] A. R. Doherty et al. "Passively recognising human activities through lifelogging". In: *Computers in Human Behavior*. Vol. 27. 5. Pergamon, 2011, pp. 1948–1958.

[19] C. Gurrin et al. "Comparing Approaches to Interactive Lifelog Search at the Lifelog Search Challenge (LSC2018)". In: *ITE Transactions on Media Technology and Applications* 7.2 (2019), pp. 46–59.

[20] A. R. Doherty and A. F. Smeaton. "Automatically Segmenting LifeLog Data into Events". In: *Proceedings of the International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*. 2008, pp. 20–23.

[21] H. Lee et al. "Constructing a SenseCam visual diary as a media process". In: *Multimedia Systems* 14.6 (2008), pp. 341–349.

[22] Y. Yang, H. Lee, and C. Gurrin. "Visualizing Lifelog Data for Different Interaction Platforms". In: *Proceedings of the Annual ACM Conference on Human Factors in Computing Systems, Extended Abstracts (CHI EA)*. Paris, France, 2013, pp. 1785–1790.

[23] Z. Qiu, C. Gurrin, and A. F. Smeaton. "Evaluating Access Mechanisms for Multimodal Representations of Lifelogs". In: *Proceedings of the International Conference on Multimedia Modeling (MMM)*. Miami, FL, USA, 2016, pp. 574–585.

[24] A. Duane, C. Gurrin, and W. Hürst. "Virtual Reality Lifelog Explorer: Lifelog Search Challenge at ACM ICMR 2018". In: *Proceedings of the ACM Workshop on the Lifelog Search Challenge (LSC@ICMR)*. Yokohama, Japan, 2018, pp. 20–23.

[25] F. Spiess et al. "Exploring Intuitive Lifelog Retrieval and Interaction Modes in Virtual Reality with vitrivr-VR". In: *Proceedings of the Annual ACM Workshop on the Lifelog Search Challenge (LSC@ICMR)*. Taipei, Taiwan, 2021, pp. 17–22.

[26] C. Gurrin et al. “Comparing Approaches to Interactive Lifelog Search at the Lifelog Search Challenge (LSC2018)”. In: *ITE Transactions on Media Technology and Applications* 7.2 (2019), pp. 46–59.

[27] A. Duane and C. Gurrin. “Lifelog Exploration Prototype in Virtual Reality”. In: *Proceedings of the International Conference on Multimedia Modeling (MMM)*. Bangkok, Thailand, 2018, pp. 377–380.

[28] H. Lee and A. F. Smeaton. “Designing Novel Applications Inspired by Emerging Media Technologies”. In: *Proceedings of the 25th BCS Conference on Human-Computer Interaction*. BCS-HCI ’11. Newcastle-upon-Tyne, United Kingdom: BCS Learning & Development Ltd., 2011, pp. 18–23.

[29] D. Norman. *Emotional design: why we love (or hate) everyday things*. Basic Books, 2004.

[30] D. Norman. “The Design of Future Things”. In: Basic Books, 2007, p. 100.

[31] E. Wilding. “The Handbook of Memory Disorders”. In: *Brain* 127.7 (2004), pp. 1681–1682.

[32] A. Sellen et al. “Do Life-Logging Technologies Support Memory for the Past? An Experimental Study Using SenseCam”. In: *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*. San Jose, CA, USA, 2007, pp. 81–90.

[33] A. R. Doherty et al. “Experiences of Aiding Autobiographical Memory Using the SenseCam”. In: *Human–Computer Interaction* 27.1-2 (2012), pp. 151–174.

[34] M. Conway. *Autobiographical memory: An introduction.* Milton Keynes, UK: Open University Press, 1990.

[35] K. Chorianopoulos. “User Interface Design Principles for Interactive Television Applications”. In: *International Journal of Human–Computer Interaction* 24.6 (2008), pp. 556–573.

[36] C. Gurrin et al. “Browsing Large Personal Multimedia Archives in a Lean-Back Environment”. In: *Proceedings of the International Conference on Multimedia Modeling (MMM)*. Chongqing, China, 2010, pp. 98–109.

[37] H. Lee et al. “Balancing the power of multimedia information retrieval and usability in designing interactive TV”. In: *uxTV 2008 - First International Conference on Designing Interactive User Experiences for TV and Video, 22-24 October 2008, Mountain View, CA, U.S.A.*. (Jan. 2008).

[38] C. Gurrin et al. “A Test Collection for Interactive Lifelog Retrieval”. In: *Proceedings of the International Conference on Multimedia Modeling (MMM)*. Thessaloniki, Greece, 2019, pp. 312–324.

[39] M. Slater and S. Wilbur. “A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments”. In: *Presence: Teleoperators and Virtual Environments* 6.6 (Dec. 1997), pp. 603–616.

[40] E. Dale. *Audio-Visual Methods in Teaching (3rd Ed.)* Dryden Press, 1969, pp. 12, 13.

[41] E. Krokos, C. Plaisant, and A. Varshney. "Virtual memory palaces: immersion aids recall". In: *Virtual Reality* 23.1 (May 2018), pp. 1–15.

[42] M. R. Mine, J. Brooks, and C. Sequin. "Moving Objects in Space : Exploiting Proprioception In Virtual-Environment Interaction". In: *24th Annual Conference on Computer Graphics and Interactive Techniques* (1997), pp. 19–26.

[43] J. Jerald. *The VR Book: Human-Centered Design for Virtual Reality*. Morgan & Claypool, 2015.

[44] A. Duane. "Visual Access to Lifelog Data in a Virtual Environment". PhD thesis. Dublin City University, Ireland: Insight Centre for Data Analytics, 2019.